Strict localization of eigenvectors and eigenvalues

Łukasz Struski* and Jacek Tabor

Jagiellonian University,
Faculty of Mathematics and Computer Science,
Łojasiewicza 6, 30-348 Kraków, Poland
e-mail: struski@ii.uj.edu.pl, tabor@ii.uj.edu.pl

Abstract

In this article we show and implement a simple and efficient method to strictly locate eigenvectors and eigenvalues of a given matrix, based on the modified cone condition. As a consequence we can also effectively localize zeros of complex polynomials.

Key words and phrases: eigenvectors, eigenvalues, cone condition, spectrum, interval arithmetic, zeros of polynomials.

2010 Mathematics Subject Classification: 65F15.

1 Introduction

1.1 Motivation

Determination of the eigenvalues and eigenvectors of a matrix is important in many areas of science, for example in web ranking [8, 15], computer graphics and visualization [24], quantum mechanics [5], statistics [12, 20], medicine, communications, construction vibration analysis [3, 23].

One of the most important numerical methods designed to calculate the eigenvalues and eigenvectors of matrix A is the power method [18, 21]. It is used to determine a maximum module eigenvalue of A and the corresponding eigenvector v. The limit of the product $\frac{A^n \mathsf{w}}{\|A^n \mathsf{w}\|}$, where w is a randomly chosen element, is the vector corresponding to the largest eigenvalue. The eigenvalue can be calculated from the Rayleigh quotient $\lambda = \frac{\langle A\mathsf{v}, \mathsf{v}\rangle}{\langle \mathsf{v}, \mathsf{v}\rangle}$. The most common method of solving the full eigenvalue problem, i.e. finding all the eigenvalues and corresponding eigenvectors, is the QR method [6, 13].

There are methods for locating eigenvalues such as Gerschgorin theorem [7] from 1931. This theorem allows to strictly locate the position of the eigenvalues of the matrix with real or complex coefficients. However, it does not allow to localize the eigenvectors.

With the growing importance of these concepts there is a need to look for new methods of localizing simultaneously eigenvalues and eigenvectors. We do not know strict and efficient methods for locating eigenvectors of real or complex matrices, so we want to fill this gap. Our aim is to create and analyze a new method of strict location eigenvectors and eigenvalues with the use of interval arithmetic[1]. Using the fact that

---

*Corresponding author
[1]For language C++ one can use libraries such as 'boost' [1] or 'CAPD' [3].

the polynomial $W(\mathsf{x}) = \mathsf{x}^n + a_{n-1}\mathsf{x}^{n-1} + \ldots + a_1\mathsf{x} + a_0$ of the n-th degree equals the determinant of the matrix

$$
\begin{bmatrix}
0 & 1 & 0 & \ldots & 0 \\
0 & 0 & 1 & \ddots & \vdots \\
\vdots & \vdots & \ddots & 1 & 0 \\
0 & 0 & \ldots & 0 & 1 \\
-a_0 & -a_1 & -a_2 & \ldots & -a_{n-1}
\end{bmatrix},
$$

we show that we can effectively localize zeros of real or complex polynomials (for non-strict methods see [10, 16]).

The content of this paper can be described as follows. In the following short subsection we present an algorithm which is the final outcome of our theory. In the Section 2 we introduce notation and present the simple properties of our theory, which is based the concept of cone condition. The notion of cone condition originally appeared in the late 60's in the works of Alekseev, Anosov, Moser and Sinai [22]. These techniques are used in the examination of differential equations [4, 9, 14, 19, 25]. In Section 3 we show the basic properties of operations. In the following section we present the main result of this article Corollary 4.1, which allow as a consequence obtain strictly localize the position of the eigenvector corresponding to the given eigenvalue in a small cone. In the last section we present a simple algorithm and some numerical applications of the theory on a few basic examples. In particular for a random matrix of size 5×5 with randomly generated values from the set $[-1, 1]$ we obtain the eigenvectors and eigenvalues with typical accuracy $\varepsilon \approx 10^{-11}$ (see Example 5.1)[2].

1.2 Basic Algorithm

In this subsection we present a basic idea how our method can be used (more detailed explanation is given in last section). This algorithm determines the accuracy for numerical approximation of the eigenvectors and the corresponding eigenvalues. Now we introduce the following notation for clarity and simplicity of the presentation.

A matrix A of dimensions $N \times N$ with elements $a_{ij} \in \mathbb{K}$ is denoted by

$$
A := [a_{ij}]_{1 \leq i,j \leq N} \in M_N(\mathbb{K}).
$$

For $K, J \subset \{1, \ldots, N\}$ by $A_{[K,J]} := [a_{kj}]_{k \in K, j \in J}$ we denote the sub-matrix with rows $k \in K$ and columns $j \in J$. We put $m : n = \{m, \ldots, n\}$ for $m, n \in \{1, \ldots, N\}$. For $k \in \{1, \ldots, N\}$ by $\neq k$ we understand the set $\{1, \ldots, N\} \setminus \{k\}$. The identity matrix is denoted by I.

By $[\![\mathsf{x}]\!]$ we understand interval representation of $\mathsf{x} \in \mathbb{R}$ that is an interval with representable ends such that $\mathsf{x} \in [\![\mathsf{x}]\!]$. For complex number $\mathsf{z} \in \mathbb{C}$ we put $[\![\mathsf{z}]\!] := [\![\mathrm{Re}(\mathsf{z})]\!] + [\![\mathrm{Im}(\mathsf{z})]\!]i$, and for a matrix $A = [a_{ij}]_{1 \leq i,j \leq N}$, $[\![A]\!]$ denotes the matrix $[\![[\![a_{ij}]\!]]\!]_{1 \leq i,j \leq N}$. We refer the reader to [17] for more information concerning interval arithmetic.

As a sup norm $\|A\|_\infty$ we take the maximum sum of absolute values of the elements in rows of matrix A, that is

$$
\|A\|_\infty = \|[a_{ij}]\|_\infty := \max_{1 \leq i \leq N} \sum_{j=1}^N \|a_{ij}\|_\infty.
$$

---

[2]In our program we use variables of type double. One could decrease the error by using numbers of arbitrary precision.

We are given a matrix $A = [a_{ij}]_{1 \leq i,j \leq N} \in M_N(\mathbb{C})$ with pairwise disjoint single eigenvalues, $\lambda_i \neq \lambda_j$ for $i \neq j$ and $i,j \in \{1,\dots,N\}$. We assume that we have calculated the numerical approximation of eigenvectors $\tilde{x}_1,\dots,\tilde{x}_N$ and eigenvalues[3] $\tilde{\lambda}_1,\dots,\tilde{\lambda}_N$ of the matrix A. To locate the exact position of k-th eigenvector x_k of A one can use the following algorithm.

Algorithm 1.1. *1)* Put $P := [\tilde{x}_1,\dots,\tilde{x}_N]$,

2) calculate interval hull $[\![P^{-1}]\!]$ of inverse of P (we call it the inverse interval matrix of P)[4],

3) calculate the interval matrix $J = [\![P^{-1}]\!] \cdot [\![A]\!] \cdot [\![P]\!]$,

4) seek a constant $\varepsilon > 0$ as small as possible so that the matrix

$$
B^\varepsilon = \begin{bmatrix} \frac{1}{\varepsilon} & 0 & \dots & 0 & & & \\ 0 & \ddots & \ddots & \vdots & & 0 & \\ \vdots & \ddots & \frac{1}{\varepsilon} & 0 & & & \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ & & 0 & \frac{1}{\varepsilon} & \ddots & \vdots & \\ & 0 & & \vdots & \ddots & \ddots & 0 \\ & & 0 & \dots & 0 & \frac{1}{\varepsilon} \end{bmatrix} \cdot (J - \tilde{\lambda}_k \cdot I) \cdot \begin{bmatrix} \varepsilon & 0 & \dots & 0 & & & \\ 0 & \ddots & \ddots & \vdots & & 0 & \\ \vdots & \ddots & \varepsilon & 0 & & & \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ & & 0 & \varepsilon & \ddots & \vdots & \\ & 0 & & \vdots & \ddots & \ddots & 0 \\ & & 0 & \dots & 0 & \varepsilon \end{bmatrix},
$$
(1.1)

where the number 1 is at the intersection of the k-th row and k-th column, satisfies the following condition

$$\|B^\varepsilon_{[k,1:N]}\|_\infty < \|(B^\varepsilon_{[\neq k, \neq k]})^{-1}\|_\infty^{-1} - \|B^\varepsilon_{[\neq k, k]}\|_\infty.$$

Then the k-th eigenvector of A satisfies

$$x_k \in [\![\tilde{x}_k]\!] + [\![-\varepsilon, \varepsilon]\!] \cdot \sum_{i \neq k}^N [\![\mathrm{Re}(\tilde{x}_i)]\!] + [\![-\varepsilon, \varepsilon]\!] \cdot \sum_{i \neq k}^N [\![\mathrm{Im}(\tilde{x}_i)]\!] i.$$

We estimate the eigenvalue λ_k from the Rayleigh coefficient

$$\lambda_k \in \frac{\langle Ax_k, x_k \rangle}{\langle x_k, x_k \rangle}.$$

We show more precise details in the last section of this paper.

Remark 1.1. Our method can be modified to work with multiple single eigenvalues (see Proposition 2.1). However, it does not work with multiple eigenvalues of the form (1.2) but this is a typical problem of most algorithms calculating the eigenvalues and eigenvectors.

$$
\begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 & 0 \\ \vdots & & \ddots & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{bmatrix},
$$
(1.2)

---

[3]To calculate the approximation of eigenvectors and eigenvalues one can use any numerical method.
[4]By $[\![P^{-1}]\!]$ we denote such an interval matrix that $Q^{-1} \in [\![P^{-1}]\!]$ for every $Q \in [\![P]\!]$.

2   Cones and Dominating Maps

In this section we modify the concept of cones from [14], which will allow us to locate the eigenvectors of the matrix.

Definition 2.1. Let E be a finite-dimensional Banach space with semi-norms $\rangle \cdot \langle$ (we call it *contracting*), $\langle \cdot \rangle$ (which we call *expanding*) such that

$$\||\mathsf{x}\|| := \max(\rangle\mathsf{x}\langle, \langle\mathsf{x}\rangle)$$

defines an equivalent norm on E. By the *r-norm* for $r > 0$ on the cone-space E we take

$$\||\mathsf{x}\||_r := \max(\rangle\mathsf{x}\langle, r \cdot \langle\mathsf{x}\rangle).$$

Definition 2.2. Let E be a cone-space. We define the *r-contracting cone* in E by

$$\rangle E\langle_r := \{\mathsf{x} \in E : \rangle\mathsf{x}\langle \geq r\langle\mathsf{x}\rangle\},$$

and the *r-expanding cone* in E by

$$\langle E\rangle_r := \{\mathsf{x} \in E : \rangle\mathsf{x}\langle \leq r\langle\mathsf{x}\rangle\}.$$

Note that

$$E = \rangle E\langle_r \cup \langle E\rangle_r. \tag{2.1}$$

In the same way we define r-contracting cone and r-expanding cone in subspace E. If $r = 1$ we will omit the subscript r, in particular we speak of contracting cone. We introduce the scaling by r of semi-norms to have better control over size of the cones (see Figure 1), and which consequently will allow us to better locate the eigenvectors.

(a) The contracting cone in $\mathbb{R} \times \mathbb{R}$.

(b) The 2-contracting cone in $\mathbb{R} \times \mathbb{R}$.

(c) The expanding cone in $\mathbb{R} \times \mathbb{R}$.

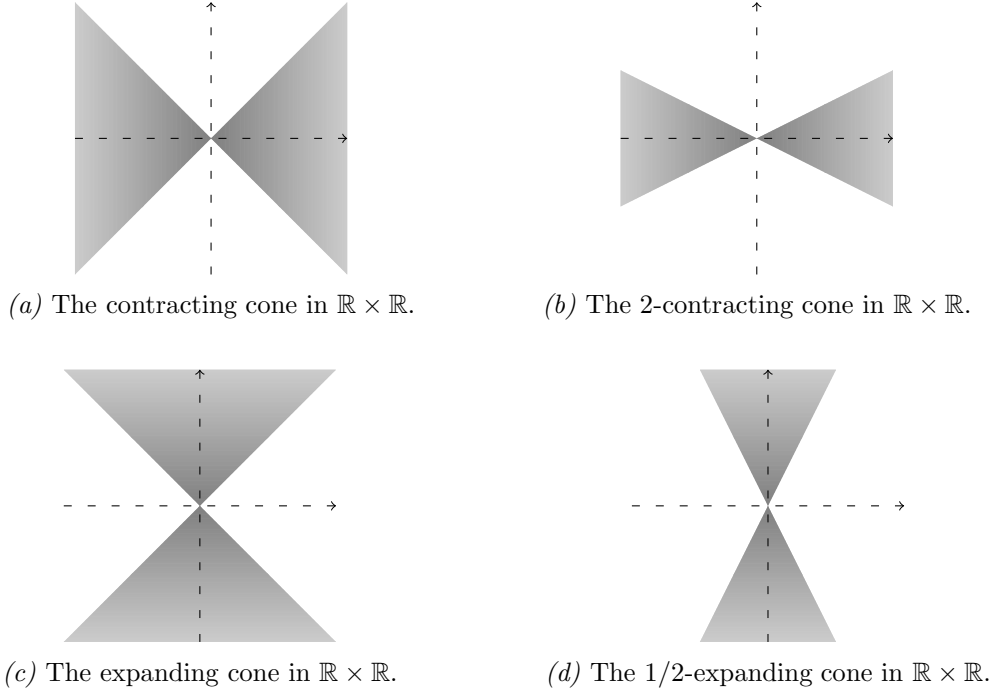*(d)* The 1/2-expanding cone in $\mathbb{R} \times \mathbb{R}$.

Figure 1: The cones in the cone-space $\mathbb{R} \times \mathbb{R}$.

For a product $E = E_1 \times E_2$ we introduce cone-space $(E, \rangle \cdot \langle, \langle \cdot \rangle)$ where we put

$$\rangle\mathsf{x}\langle := \|x_1\|, \quad \langle\mathsf{x}\rangle := \|x_2\| \quad \text{for } \mathsf{x} = (x_1, x_2) \in E_1 \times E_2.$$

Analogically, we define cone-space for direct sum $E = E_1 \oplus E_2$.

In our main result, Corollary 4.1, the following proposition will play a crucial role.

4

Proposition 2.1. *Let $E = E_1 \times E_2$ be cone-space such that $\dim E_1 = n$, $\dim E_2 = m$ and let $r > 0$ be given. Assume that we have direct sum decomposition $E = V_1 \oplus V_2$ such that*

$$V_1 \subset \rangle E\langle_r \quad and \quad V_2 \subset \langle E\rangle_r.$$

Then

$$\dim V_1 = n \quad and \quad \dim V_2 = m.$$

Proof. First we show that $\dim V_1 \leq n$. For an indirect proof, assume that $\dim V_1 > n$. Then there exist linearly independent vectors $\mathsf{v}_1, \ldots, \mathsf{v}_{\mathsf{n+1}} \in V_1$. Obviously $\mathsf{v}_i = (w_i, z_i)$ for $i \in \{1, \ldots, n+1\}$ and uniques $w_i \in E_1$, $z_i \in E_2$. Since $w_1, \ldots, w_{n+1} \in E_1$ and $\dim E_1 = n$ there exist a set of $n+1$ scalars, $\alpha_1, \ldots, \alpha_{n+1}$, not all zero, such that

$$\alpha_1 w_1 + \ldots + \alpha_{n+1} w_{n+1} = 0.$$

Note that

$$z := \alpha_1 z_1 + \ldots + \alpha_{n+1} z_{n+1} \neq 0,$$

because otherwise vectors $\mathsf{v}_1, \ldots, \mathsf{v}_{\mathsf{n+1}}$ would not be linearly independent. Consequently we obtain

$$(0, z) = \left(\sum_{i=1}^{n+1} \alpha_i w_i, \sum_{i=1}^{n+1} \alpha_i z_i \right) \in V_1 \subset \rangle E\langle_r,$$

and thus $r\|z\| \leq \|0\|$, which implicate $z = 0$. We get a contradiction with the fact the sequence of vectors $\mathsf{v}_1, \ldots, \mathsf{v}_{\mathsf{n+1}}$ is linearly independent.

The proof that $\dim V_2 \leq m$ is analogous. Finally, since $\dim E = n + m$ and $\dim V_1 \leq n$, $\dim V_2 \leq m$ we obtain

$$\dim V_1 = n, \quad and \quad \dim V_2 = m.$$

$\square$

By an *operator* we mean a linear mapping between cone-spaces E and F. We denote the space of all operators by $\mathcal{L}(E, F)$. If $F = E$, we denote $\mathcal{L}(E, E)$ by $\mathcal{L}(E)$.

Let $A \in \mathcal{L}(E, F)$. We define

$$\rangle A\langle_r := \inf\{R \in \mathbb{R}_+ \mid \||A\mathsf{x}\||_r \leqslant R \||\mathsf{x}\||_r \text{ for all } \mathsf{x} \in E : A\mathsf{x} \in \rangle F\langle_r\}, \qquad (2.2)$$

$$\langle A\rangle_r := \sup\{R \in \mathbb{R}_+ \mid \||A\mathsf{x}\||_r \geqslant R \||\mathsf{x}\||_r \text{ for all } \mathsf{x} \in E : \mathsf{x} \in \langle E\rangle_r\}. \qquad (2.3)$$

Let $\tilde{E} \subset E$, $\tilde{F} \subset F$ be subspaces and let $A \in \mathcal{L}(E, F)$ such that $A(\tilde{E}) \subset \tilde{F}$. We define

$$\rangle A|_{\tilde{E}}\langle_r := \inf\{R \in \mathbb{R}_+ \mid \||A\mathsf{x}\||_r \leqslant R \||\mathsf{x}\||_r \text{ for all } \mathsf{x} \in \tilde{E} : A\mathsf{x} \in \rangle\tilde{F}\langle_r\},$$

$$\langle A|_{\tilde{E}}\rangle_r := \sup\{R \in \mathbb{R}_+ \mid \||A\mathsf{x}\||_r \geqslant R \||\mathsf{x}\||_r \text{ for all } \mathsf{x} \in \tilde{E} : \mathsf{x} \in \langle\tilde{E}\rangle_r\}.$$

Definition 2.3. We say that $A \in \mathcal{L}(E, F)$ is *r-dominating*, if

$$\rangle A\langle_r < \langle A\rangle_r.$$

By $\mathcal{D}_r(E, F)$ we denote the set of all $A \in \mathcal{L}(E, F)$ which are r-dominating. If $F = E$, we denote the space $\mathcal{D}_r(E, E)$ by $\mathcal{D}_r(E)$.

Observation 2.1. *Let $\tilde{E} \subset E$, $\tilde{F} \subset F$ be subspaces and let $A \in \mathcal{L}(E, F)$ such that $A(\tilde{E}) \subset \tilde{F}$. We have*

$$\rangle A|_{\tilde{E}}\langle_r \leq \rangle A\langle_r \quad and \quad \langle A\rangle_r \leq \langle A|_{\tilde{E}}\rangle_r.$$

Moreover, if $A \in \mathcal{D}_r(E, F)$ then $A \in \mathcal{D}_r(\tilde{E}, \tilde{F})$.

Proof. It is consequence of (2.2), (2.3) and Definition 2.3. □

Theorem 2.1. *Let $A \in \mathcal{D}_r(E, F)$ and let $\mathsf{v} \in E$ be arbitrary. Then*

$$\mathsf{v} \in \langle E \rangle_r \implies A\mathsf{v} \in \langle F \rangle_r,$$
$$A\mathsf{v} \in \rangle F \langle_r \implies \mathsf{v} \in \rangle E \langle_r.$$

Proof. The proof is a simple modification of the proof of [14, Proposition 2.1]. □

Proposition 2.2. *Let $A \in \mathcal{D}_r(F, G)$ and $B \in \mathcal{D}_r(E, F)$. Then $A \circ B \in \mathcal{D}_r(E, G)$ and*

$$\rangle A \circ B \langle_r \le \rangle A \langle_r \cdot \rangle B \langle_r, \quad \langle A \circ B \rangle_r \ge \langle A \rangle_r \cdot \langle B \rangle_r. \tag{2.4}$$

Proof. To prove the first inequality from (2.4), let $\mathsf{x} \in E$ and $A \circ B(\mathsf{x}) \in \rangle G \langle_r$. From (2.2) and Theorem 2.1 we know that $B\mathsf{x} \in \rangle F \langle_r$. We have

$$|\!|\!| A \circ B(\mathsf{x}) |\!|\!|_r \le \rangle A \langle \, |\!|\!| B\mathsf{x} |\!|\!|_r \le \rangle A \langle \, \rangle B \langle \, |\!|\!| \mathsf{x} |\!|\!|_r.$$

Hence

$$\rangle A \circ B \langle \le \rangle A \langle \cdot \rangle B \langle.$$

Using (2.3) and Theorem 2.1, we obtain the second inequality from (2.4).

As a simple consequence of (2.4) we obtain $A \circ B \in \mathcal{D}_r(E, G)$. □

3 Operator Norms

In this section we show how given an operator A to estimate $\rangle A \langle_r$, $\langle A \rangle_r$.

Consider two cone-spaces $E = E_1 \times E_2$ and $F = F_1 \times F_2$. Let $A \colon E \to F$ be an operator given in the matrix form by

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}.$$

By

$$|\!|\!| A |\!|\!|_r := \max \left(\|A_{11}\| + \frac{1}{r}\|A_{12}\|, r\|A_{21}\| + \|A_{22}\| \right) \tag{3.1}$$

we define the r-norm of operator A. Observe that it satisfies

$$|\!|\!| A\mathsf{x} |\!|\!|_r \le |\!|\!| A |\!|\!|_r \cdot |\!|\!| \mathsf{x} |\!|\!|_r \quad \text{for } \mathsf{x} \in E.$$

Note that in general it is not, if E_1 is not one dimensional, the classical operator norm for $|\!|\!| \cdot |\!|\!|_r$.

Observation 3.1. *Let $r \in (0, \infty)$, $A \in \mathcal{L}(E_1 \times E_2, F_1 \times F_2)$ We put*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad and \quad R = \begin{bmatrix} I & 0 \\ 0 & rI \end{bmatrix}.$$

Then

$$|\!|\!| \mathsf{x} |\!|\!|_r = |\!|\!| R\mathsf{x} |\!|\!| \quad for \ \mathsf{x} \in E,$$
$$|\!|\!| A |\!|\!|_r = |\!|\!| RAR^{-1} |\!|\!|.$$

Theorem 3.1. *Let $A = [A_{ij}]_{1 \le i, j \le 2} \in \mathcal{L}_r(E_1 \times E_2, F_1 \times F_2)$.*

i) We have

$$\rangle A\langle_r \leq \|A_{11}\| + \frac{1}{r}\|A_{12}\|.$$

ii) Additionally, if A_{22} is invertible, then

$$\langle A\rangle_r \geq \|A_{22}^{-1}\|^{-1} - r\|A_{21}\|.$$

Proof. We show the proof for $r = 1$. For arbitrary $r \in (0, \infty)$ the proof needs a simple modification, see Observation 3.1.

For the proof of the first inequality, we take $\mathsf{x} = (x_1, x_2) \in E_1 \times E_2$ such that $A\mathsf{x} \in \rangle F\langle$. From Definition 2.2 we have

$$\|A_{11}x_1 + A_{12}x_2\| \geq \|A_{21}x_1 + A_{22}x_2\|, \tag{3.2}$$

therefore

$$\||A\mathsf{x}|\| = \max(\|A_{11}x_1 + A_{12}x_2\|, \|A_{21}x_1 + A_{22}x_2\|)$$
$$\overset{(3.2)}{=} \|A_{11}x_1 + A_{12}x_2\| \leq \{\|A_{11}\| \cdot \|x_1\| + \|A_{12}\| \cdot \|x_2\|\}$$
$$\leq (\|A_{11}\| + \|A_{12}\|) \cdot \||\mathsf{x}|\|.$$

For the proof of the second inequality, suppose that $\mathsf{x} = (x_1, x_2) \in \langle E\rangle$, where $x_1 \in E_1$, $x_2 \in E_2$. Then

$$\|x_1\| \leq \|x_2\| = \||\mathsf{x}|\|. \tag{3.3}$$

We know that

$$\|A_{22}x_2\| \geq \|A_{22}^{-1}\|^{-1}\|x_2\| \geq 0. \tag{3.4}$$

Finally, we obtain

$$\||A\mathsf{x}|\| \geq \|A_{21}x_1 + A_{22}x_2\| \geq \|A_{22}x_2\| - \|A_{21}x_1\|$$
$$\overset{(3.4)}{\geq} \|A_{22}^{-1}\|^{-1}\|x_2\| - \|A_{21}\|\|x_1\| \overset{(3.3)}{\geq} \left(\|A_{22}^{-1}\|^{-1} - \|A_{21}\|\right) \cdot \||\mathsf{x}|\|.$$

$\square$

Example 3.1. Let us verify that the matrix $A \in \mathcal{L}(\mathbb{R} \times \mathbb{R}, \mathbb{R} \times \mathbb{R})$

$$A = \begin{bmatrix} 2 & 1.5 \\ 1 & 5 \end{bmatrix}$$

is dominating.

By Theorem 3.1 we have

$$\rangle A\langle \leq 3.5 < 4 \leq \langle A\rangle,$$

and therefore A is dominating.

4   The Main Results

In this section we show the main results of the paper, concerns the strict location of the eigenspace.

We denote the spectrum of the operator A by $\sigma(A) := \{\lambda \in \mathbb{C} \ : \ A - \lambda I \text{ is singular}\}$.

Lemma 4.1. *Let $A \in \mathcal{D}_r(E)$. Then*

$$\lambda \in \sigma(A) \implies |\lambda| \subset [0, \,\rangle A\langle_r\,] \cup [\langle A\rangle_r, \infty). \tag{4.1}$$

Moreover $[0, \,\rangle A\langle_r\,] \cap [\langle A\rangle_r, \infty) = \emptyset$.

Proof. Since $A \in \mathcal{D}_r(E)$ we get $[0, \,\rangle A\langle_r\,] \cap [\langle A\rangle_r, \infty) = \emptyset$.

Now we show implication (4.1). Let λ be an eigenvalue of A and let $\mathsf{x} \in E$ be a corresponding eigenvector. By (2.1) we know that $\mathsf{x} \in \,\rangle E\langle_r \cup \langle E\rangle_r$. We consider two cases. First suppose that $\mathsf{x} \in \,\rangle E\langle_r$. Since x is an eigenvector, $A\mathsf{x} = \lambda\mathsf{x}$, and thus $A\mathsf{x} \in \,\rangle E\langle_r$. By (2.2) we get

$$|\lambda| \leq \,\rangle A\langle_r .$$

Now suppose that $\mathsf{x} \in \langle E\rangle_r$. By (2.3) we get

$$|\lambda| \geq \langle A\rangle_r,$$

which complete the proof. $\qquad\square$

Let E be a vector space over the field $\mathbb{K}$ and let operator $A\colon E \to E$ be given. One can easy deduce from the theorem of Jordan, see [11, Appendix to Chapter 4] for the general case, that if $\sigma = \sigma_1 \cup \sigma_2$ (for $\mathbb{K} = \mathbb{R}$ we assume additional that $\bar{\sigma}_1 = \sigma_1$ and $\bar{\sigma}_2 = \sigma_2$) then there is a unique direct sum decomposition $E = E_{\sigma_1} \oplus E_{\sigma_2}$ such that $A(E_{\sigma_1}) \subset E_{\sigma_1}$, $A(E_{\sigma_2}) \subset E_{\sigma_2}$ and $\sigma(A|_{E_{\sigma_1}}) = \sigma_1$, $\sigma(A|_{E_{\sigma_2}}) = \sigma_2$. For $c > 0$ we define

$$E_{\leq c} := E_{\{\lambda \,:\, |\lambda| \leq c\}} \quad \text{and} \quad E_{\geq c} := E_{\{\lambda \,:\, |\lambda| \geq c\}}.$$

Theorem 4.1. *Let E be a finite-dimensional cone-space and let $A \in \mathcal{D}_r(E)$. Then there is a direct sum decomposition*

$$E = E_{\leq \,\rangle A\langle_r} \oplus E_{\geq \langle A\rangle_r} \quad \text{and} \quad E_{\leq \,\rangle A\langle_r} \subset \,\rangle E\langle_r\,, \ E_{\geq \langle A\rangle_r} \subset \langle E\rangle_r.$$

Proof. For the clarity of presentation we consider only the case $r = 1$ and we omit the subscript r.

From Lemma 4.1 and Jordan Theorem we obtain a direct sum decomposition

$$E = E_{\leq \,\rangle A\langle} \oplus E_{\geq \langle A\rangle},$$

such that

$$\sigma(A|_{E_{\leq \,\rangle A\langle}}) = [0, \,\rangle A\langle\,] \quad \text{and} \quad \sigma(A|_{E_{\geq \langle A\rangle}}) = [\langle A\rangle, \infty).$$

Now we show $E_{\leq \,\rangle A\langle} \subset \,\rangle E\langle$. Consider an arbitrary $\mathsf{x} \in E_{\leq \,\rangle A\langle}$. The case when $\mathsf{x} = 0$ is obvious. Assume that $\mathsf{x} \neq 0$. Without loss of generality we can assume that $\|\mathsf{x}\| = 1$. For an indirect proof, assume that $\mathsf{x} \notin \,\rangle E\langle$. Then $\mathsf{x} \in \langle E\rangle$ (by 2.1). Let $\varepsilon > 0$ be arbitrary. From the fact that $\mathsf{x} \in E_{\leq \,\rangle A\langle}$, we know that

$$\limsup_{n \to +\infty} \sqrt[n]{\||A^n\mathsf{x}\||} \leq \,\rangle A\langle\,,$$

and thus there exists $m \in \mathbb{N}$ such that for all $n \in \mathbb{N}$

$$n \geq m \Rightarrow \sqrt[n]{\||A^n\mathsf{x}\||} \leq \,\rangle A\langle + \varepsilon.$$

Since $\mathsf{x} \in \langle E\rangle$ and from Theorem 2.1 we obtain

$$\mathsf{x} \in \langle E\rangle \Rightarrow A\mathsf{x} \in \langle E\rangle \Rightarrow \cdots \Rightarrow A^n\mathsf{x} \in \langle E\rangle.$$

Using (2.3) we get

$$\||A\mathsf{x}\|| \geq \langle A \rangle \, \||\mathsf{x}\||,$$
$$\||A^2\mathsf{x}\|| = \||A(A\mathsf{x})\|| \geq \langle A \rangle \, \||\,A\mathsf{x}\|| \geq \langle A \rangle^2 \, \||\mathsf{x}\||,$$
$$\vdots$$
$$\||A^n\mathsf{x}\|| \geq \langle A \rangle^n \, \||\,\mathsf{x}\,\||\,.$$

Finally we have

$$\langle A \rangle = \sqrt[n]{\langle A \rangle^n} \leq \sqrt[n]{\||A^n\mathsf{x}\||} \leq \rangle A \langle + \varepsilon.$$

Since ε was arbitrary, we get a contradiction with the fact that A is r-dominating.

Analogously to the proof of the second conclusion, assume that $\mathsf{x} \in E_{\geq \langle A \rangle}$ and $\mathsf{x} \notin \langle E \rangle$. Then $\mathsf{x} \in \rangle E \langle$. Since $\sigma(A|_{E_{\geq \langle A \rangle}}) = \sigma_{\geq \langle A \rangle} := \{\lambda \ : \ |\lambda| \geq \langle A \rangle\}$ and $0 \notin \sigma_{\geq \langle A \rangle}$ we know that $A|_{E_{\geq \langle A \rangle}} : E_{\geq \langle A \rangle} \to E_{\geq \langle A \rangle}$ is invertible. Let $\varepsilon > 0$ be arbitrary. Using the fact that $\mathsf{x} \in E_{\geq \langle A \rangle}$ we know that

$$\limsup_{n \to +\infty} \sqrt[n]{\||A|_{E_{\geq \langle A \rangle}}^{-n}\mathsf{x}\||} \leq \langle A \rangle^{-1},$$

and thus there exists $m \in \mathbb{N}$ such that for all $n \in \mathbb{N}$

$$n \geq m \Rightarrow \sqrt[n]{\||A|_{E_{\geq \langle A \rangle}}^{-n}\mathsf{x}\||} \leq \langle A \rangle^{-1} + \varepsilon. \tag{4.2}$$

From Observation 2.1 and Theorem 2.1 we get

$$\mathsf{x} \in \rangle E_{\geq \langle A \rangle}\langle_r \Rightarrow A|_{E_{\geq \langle A \rangle}}^{-1}\mathsf{x} \in \rangle E_{\geq \langle A \rangle}\langle \Rightarrow \cdots \Rightarrow A|_{E_{\geq \langle A \rangle}}^{-n}\mathsf{x} \in \rangle E_{\geq \langle A \rangle}\langle\,,$$

and from (2.2) we have

$$\||\mathsf{x}\|| \leq \rangle A|_{E_{\geq \langle A \rangle}}\langle \, \||A|_{E_{\geq \langle A \rangle}}^{-1}\mathsf{x}\||,$$
$$\||A|_{E_{\geq \langle A \rangle}}^{-1}\mathsf{x}\|| \leq \rangle A|_{E_{\geq \langle A \rangle}}\langle \, \||A|_{E_{\geq \langle A \rangle}}^{-2}\mathsf{x}\||,$$
$$\vdots$$
$$\||A|_{E_{\geq \langle A \rangle}}^{-n+1}\mathsf{x}\|| \leq \rangle A|_{E_{\geq \langle A \rangle}}\langle \, \||A|_{E_{\geq \langle A \rangle}}^{-n}\mathsf{x}\,\||\,.$$

Hence

$$\||\,\mathsf{x}\|| \leq (\rangle A|_{E_{\geq \langle A \rangle}}\langle\,)^n \, \||\,A|_{E_{\geq \langle A \rangle}}^{-n}\mathsf{x}\,\||\,. \tag{4.3}$$

Finally from Observation 2.1 and (4.2), (4.3) we obtain

$$\rangle A \langle \geq \rangle A|_{E_{\geq \langle A \rangle}}\langle_r = \sqrt[n]{(\rangle A|_{E_{\geq \langle A \rangle}}\langle\,)^n} \geq \sqrt[n]{\frac{1}{\||A|_{E_{\geq \langle A \rangle}}^{-n}\mathsf{x}\||}} \geq \frac{1}{\langle A \rangle^{-1} + \varepsilon} = \langle A \rangle \cdot \frac{1}{1 + \varepsilon \cdot \langle A \rangle},$$

which get a contradiction with the fact that A is r-dominating. $\qquad\square$

By $\mathbb{I}$ we denote interval $[\![-1, 1]\!]$. For $\varepsilon > 0$ we put $B_{\mathbb{C}}(0, \varepsilon) := \{\mathsf{z} \in \mathbb{C} \ : \ |\mathsf{z}| \leq \varepsilon\}$.

Corollary 4.1. Let $\varepsilon > 0$ and $N \in \mathbb{N}$. Assume that an operator $A \in \mathcal{D}_r(\mathbb{K} \times \mathbb{K}^{N-1})$ for $r = 1/\varepsilon$ is given.

  i) $\mathbb{K} = \mathbb{R}$. Then there exist unique eigenvalue λ of A such that $|\lambda| \leq \rangle A \langle_r$ and the eigenspace is one-dimensional space. The unique (module rescaling) eigenvector x corresponding to the eigenvalue λ satisfies

$$\mathsf{x} \in (1, 0, \ldots, 0)^T + \varepsilon \cdot (0, \mathbb{I}, \ldots, \mathbb{I})^T.$$

ii) $\mathbb{K} = \mathbb{C}$. Then there exist unique eigenvalue λ of A such that $|\lambda| \leq \rangle A\langle_r$ and the eigenspace is one-dimensional space. The unique (module rescaling) eigenvector x corresponding to the eigenvalue λ satisfies

$$\mathsf{x} \in (1,0,\ldots,0)^T + \{0\} \times B_{\mathbb{C}}(0,\varepsilon)^{N-1} \subset (1,0,\ldots,0)^T + \varepsilon \cdot (0,\mathbb{I},\ldots,\mathbb{I})^T + \varepsilon \cdot (0,\mathbb{I},\ldots,\mathbb{I})^T i.$$

Proof. It is a direct consequence of Theorem 4.1 and Proposition 2.1. $\qquad\square$

Example 4.1. Consider the following square matrix

$$A = \begin{bmatrix} 1 & 0.4 & 0.5 \\ 0.4 & 4 & 0.4 \\ 0.5 & 0.4 & 8 \end{bmatrix}.$$

We find the localization of x_1 eigenvector corresponding to the eigenvalue of smallest module. Since A satisfies for $r = 1.1$

$$\rangle A\langle_r \leq 1.81(81) < 3.24 \leq \langle A\rangle_r,$$

we can use Theorem 4.1. We get

$$\mathsf{x}_1 \in \{(x_1, x_2, x_3) \; : \; \max\{|x_2|, |x_3|\} \leq \frac{10}{11}|x_1|\}. \tag{4.4}$$

Finding the roots of the characteristic polynomial of A, which is equal to

$$W(\lambda) = -\lambda^3 + 13\lambda^2 - 43.43\lambda + 29.72$$

$$= -(\lambda - 4)(\lambda - \frac{9 - \sqrt{51.28}}{2})(\lambda - \frac{9 + \sqrt{51.28}}{2}).$$

we get

$$\mathsf{x}_1 = \left(\frac{\lambda_1 - 7.5}{\lambda_1 - 0.5}x, \frac{0.8\lambda_1 - 3.2}{(\lambda_1 - 4)(\lambda_1 - 0.5)}x, x \right) \quad \text{for} \;\; x \in \mathbb{R},$$

$$\mathsf{x}_1 \approx (-15.686641x, 1.9070447x, x) \qquad \text{for} \;\; x \in \mathbb{R},$$

Note that x_1 satisfies (4.4).

5 Applications

In this section we show the numerical applications of our theory. For the purposes of this article, we use the programming language c++ with 'Boost' libraries [1] (version 1.51.0), which contains the interval arithmetic, and software library for numerical linear algebra 'Linear Algebra PACKage' [2] (revision 1348). Alternatively, one could use numerical computing environment 'Matlab' with an interval package. In the first section we presented a simple algorithm (see Algorithm 1.1), which strictly localizes the position of particular eigenvalues and eigenvectors of a matrix.

Now we discuss in more details the steps from the Algorithm 1.1. First we determine the numerical approximation of eigenvectors $\tilde{\mathsf{x}}_1, \ldots, \tilde{\mathsf{x}}_N$ and corresponding eigenvalues $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_N$ of the matrix $A \colon \mathbb{C}^N \to \mathbb{C}^N$. In the step 1) of this algorithm we create matrix P of numerical approximate eigenvectors $\tilde{\mathsf{x}}_1, \ldots, \tilde{\mathsf{x}}_N$. For the step 2) we calculate the inverse matrix of the interval matrix which is composed of the approximate eigenvectors.

The inverse matrix we can determine by the method of Gauss-Jordan elimination. In the next step we calculate the interval matrix

$$J = [\![P^{-1}]\!] \cdot [\![A]\!] \cdot [\![P]\!].$$

The most important stage of our method is the step 4). In this step we want to take advantage of Corollary 4.1, but first we need to check whether our matrix J satisfies the assumptions of this corollary (is it r-dominating). We want to localize the k-th eigenvector of a general matrix. We create matrices B^ε (see (1.1)) and we check whether the following inequality is satisfied (see Theorem 3.1):

$$\rangle B^\varepsilon \langle \; < \; \langle B^\varepsilon \rangle.$$

If we find an ε which satisfies the above condition for the matrix B^ε, then from Corollary 4.1 we know that the k-th eigenvector of J satisfies

$$(1, 0, \ldots, 0)^T + \varepsilon \cdot (0, \mathbb{I}, \ldots, \mathbb{I})^T + \varepsilon \cdot (0, \mathbb{I}, \ldots, \mathbb{I})^T i,$$

where $\mathbb{I} = [\![-1, 1]\!]$. Since $J = [\![P^{-1}]\!] \cdot [\![A]\!] \cdot [\![P]\!]$ we obtain the k-th eigenvector of A multiplying the k-th eigenvector of J by interval matrix $[\![P]\!]$. Finally we have

$$\mathsf{x}_\mathsf{k} \in [\![\tilde{\mathsf{x}}_\mathsf{k}]\!] + [\![-\varepsilon, \varepsilon]\!] \cdot \sum_{i \neq k}^{N} [\![\mathrm{Re}(\tilde{\mathsf{x}}_\mathsf{i})]\!] + [\![-\varepsilon, \varepsilon]\!] \cdot \sum_{i \neq k}^{N} [\![\mathrm{Im}(\tilde{\mathsf{x}}_\mathsf{i})]\!] i.$$

The eigenvalue λ_k corresponding to the eigenvector x_k we estimate from

$$\lambda_k \in \frac{\langle A\mathsf{x}_\mathsf{k}, \mathsf{x}_\mathsf{k} \rangle}{\langle \mathsf{x}_\mathsf{k}, \mathsf{x}_\mathsf{k} \rangle}. \tag{5.1}$$

Now we show some examples where we used our method.

Example 5.1. Consider 500 random matrices of size 5×5 with randomly generated values from the set $[-1, 1]$. For each such matrix we launched our program to find accurate ε location of all eigenvectors. We obtained $\varepsilon \in [10^{-15}, 10^{-5}]$ and for the results we have the fist-quantile to equal $6 \cdot 10^{-13}$, the median to equal 10^{-11} and the third-quantile to equal $8 \cdot 10^{-10}$.

Example 5.2. Let A be given in the matrix form

$$A = \begin{bmatrix}
-2 & 0 & 3 & 1 & -1 & 3 & -3 & -2 & 4 & 1 \\
0 & 3 & -1 & 4 & 2 & 2 & -3 & 0 & -4 & 0 \\
0 & 0 & -3 & -4 & 2 & -2 & 0 & -3 & -1 & 1 \\
-4 & -1 & 4 & 1 & -1 & 2 & 4 & 1 & 2 & 0 \\
3 & -1 & 4 & 0 & 4 & 3 & -2 & 0 & 1 & 3 \\
4 & -1 & 1 & 2 & 1 & -4 & 2 & -2 & -4 & -2 \\
0 & 4 & 1 & -1 & -2 & -4 & -2 & 4 & 1 & -1 \\
-3 & -4 & 2 & 0 & 0 & 0 & 0 & 0 & 4 & 0 \\
4 & -4 & 2 & 0 & -1 & 0 & -2 & -4 & 4 & 0 \\
-4 & -3 & 4 & 4 & 0 & 4 & -3 & 3 & 3 & -1
\end{bmatrix}.$$

Our program found the following set of $\varepsilon \in [10^{-10}, 10^{-8}]$.

Due to the large size of eigenvectors of A we write only first of them

$$
\begin{aligned}
\mathsf{x}_1 \in (\ & 0.0032301 + [32, 51] \cdot 10^{-9} + (-0.1909691 + [-14, -10] \cdot 10^{-9})i, \\
& 0.0032301 + [20, 62] \cdot 10^{-9} + (0.1909691 + [9, 16] \cdot 10^{-9})i, \\
& 0.1312400 + [61, 85] \cdot 10^{-9} + [-34, 34] \cdot 10^{-10}i, \\
& 0.2755005 + [42, 60] \cdot 10^{-9} + [-28, 28] \cdot 10^{-10}i, \\
& 0.1849394 + [66, 98] \cdot 10^{-9} + (-0.0068575 + [-54, -45] \cdot 10^{-9})i, \\
& 0.1849394 + [66, 98] \cdot 10^{-9} + (0.0068575 + [45, 54] \cdot 10^{-9})i, \\
& 0.3754802 + [56, 77] \cdot 10^{-9} + (0.0266195 + [36, 44] \cdot 10^{-9})i, \\
& 0.3754802 + [56, 77] \cdot 10^{-9} + (-0.0266195 + [-43, -36] \cdot 10^{-9})i, \\
& 0.1456179 + [89, 94] \cdot 10^{-9} + [-54, 54] \cdot 10^{-11}i, \\
& -0.0358730 + [-7, 0] \cdot 10^{-9} + [-78, 78] \cdot 10^{-11}i),
\end{aligned}
$$

The eigenvalue corresponding to the eigenvector x_1 satisfies

$$
\lambda_1 \in 5.56625 + [46, 81] \cdot 10^{-7} + (3.1629 + [69, 72] \cdot 10^{-6})i.
$$

Example 5.3. Consider the polynomial

$$
W(\mathsf{x}) = \mathsf{x}^5 + (5 - i)\mathsf{x}^4 - 7i\mathsf{x}^2 + (2 + 4i)\mathsf{x} - 8.
$$

To find all the roots of polynomial $W(\mathsf{x})$ we estimate the eigenvalues of the matrix

$$
A = \begin{bmatrix}
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 \\
8 & -2 - 4i & 7i & 0 & -5 + i
\end{bmatrix}.
$$

From our program we obtain

$$
\begin{aligned}
\lambda_1 &\in -5.1189735 + [-66, -07] \cdot 10^{-9} + 1.2610393 + [39, 67] \cdot 10^{-9}i, \\
\lambda_2 &\in 0.384 + [37, 42] \cdot 10^{-5} + 1.2215 + [45, 97] \cdot 10^{-6}i, \\
\lambda_3 &\in 0.9572 + [53, 90] \cdot 10^{-6} + 0.1374 + [16, 42] \cdot 10^{-6}i, \\
\lambda_4 &\in -1.0805 + [-80, -10] \cdot 10^{-6} - 0.6647 + [-95, -42] \cdot 10^{-6}i, \\
\lambda_5 &\in -0.1421 + [-68, -30] \cdot 10^{-6} - 0.9552 + [-97, -45] \cdot 10^{-6}i.
\end{aligned}
$$

References

[1] Boost C++ Libraries. Website: www.boost.org.

[2] LAPACK – Linear Algebra PACKage. Website: www.netlib.org/lapack.

[3] The CAPD library. Website: capd.ii.uj.edu.pl.

[4] M. J. Capiński and P. Zgliczyński. Cone conditions and covering relations for topologically normally hyperbolic invariant manifolds. Discrete and Continuous Dynamical Systems - Series A, 30:641–670, 2011.

[5] C. Cohen-Tannoudji. Chapter II. The mathematical tools of quantum mechanics in: Quantum mechanics. New York : Wiley, 1977.

[6] J.G.F. Francis. The QR transformation I. The Computer Journal, 4(3):265–271, 1961.

[7] S. Gerschgorin. Über die Abgrenzung der Eigenwerte einer Matrix. Bulletin de l'Académie des Sciences de l'URSS. Classe des sciences mathématiques et na, 6:749–754, 1931.

[8] T. Haveliwala, S. Kamvar, and G. Jeh. An Analytical Comparison of Approaches to Personalizing PageRank. Stanford University Technical Report, 2003.

[9] S. Hruska, *A Numerical Method for Constructing the Hyperbolic Structure of Comple Hénon Mappings*, Found. Comput. Math., 6(4):427–455, 2006.

[10] J. Hubbard, D. Schleicher, and S. Sutherland. How to find all roots of complex polynomials by Newtons method. Inventiones Mathematicae, 146:1–33, 2001.

[11] M. C. Irwin. Smooth dynamical systems. Advanced Series in Nonlinear Dynamics, 17. World Scientific Publishing Co., Inc., River Edge, NJ, 2001. xii+259 pp. ISBN: 981-02-4599-8, Reprint of the 1980 original. With a foreword by R. S. MacKay.

[12] I. T. Jolliffe. Principal Component Analysis. Springer-Verlag, page 487, 1986.

[13] V. N. Kublanovskaya. On some algorithms for the solution of the complete eigenvalue problem. USSR Computational Mathematics and Mathematical Physics, 1:637–657, 1962.

[14] T. Kułaga and J. Tabor. Hyperbolic dynamics in graph-directed IFS. Journal of Differential Equations, 251(12):3363–3380, 2011.

[15] A. N. Langvillei and C. D. Meyer, Google's PageRank and Beyond: The Science of Search Engine Rankings. The Mathematical Intelligencer, 30(1):68–69, 2008.

[16] J.M. McNamee. Numerical methods for roots of polynomials. Elsevier Science, 2007.

[17] Ramon E. Moore, R. Baker Kearfott, and Michael J. Cloud. Introduction to interval analysis. pub-SIAM, 2009.

[18] R. B. Morgan. Computing interior eigenvalues of large matrices. Linear Algebra and its Applications, 154–156:289–309, 1991.

[19] Sh. Newhouse. Cone-fields, Domination, and Hyperbolicity. Modern Dynamical Systems and Applications, pages 419–433, 2004.

[20] K. Pearson. On Lines and Planes of Closest Fit to Systems of Points in Space, volume 2 of 6. Philosophical Magazine, 1901.

[21] Gerard L. G. Sleijpen and Henk A. Van der Vorst. A Jacobi-Davidson Iteration Method for Linear Eigenvalue Problems. SIAM Review, 42(2):267–293, 2000.

[22] R. Sturman, J. M. Ottino, and S. Wiggins. The mathematical foundations of mixing. Cambridge University Press., 2006.

[23] Tustin Wayne. Where to place the control accelerometer. EE-Evaluation Engineering, 10(45):50, 2006.

[24] Xirouhakis Y., Votsis G. i Delopoulos, A., Estimation of 3D Motion and Structure of Human Faces. In Advances in intelligent systems : concepts, tools and applications, 1998.

[25] P. Zgliczyński. Covering relations, cone conditions and stable manifold theorem. Journal of Differential Equations, 246:1774–1819, 2009.